DOCUMENT RESUME

ED 171 038 FL 010 329

AUTHOR Konttinen, Raimo
TITLE An English Language Teaching Experiment in the Comprehension of Scientific Texts. Further Contrastive Papers, Jyvaskyla Contrastive Studies, 6. Reports from the Department of English, No. 7.
INSTITUTION Jyvaskyla Univ. (Finland). Dept. of English.
PUB DATE 78
NOTE 25p.
AVAILABLE FROM Department of English, University of Jyvaskyla, 40100 Jyvaskyla 10, Finland

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Objectives; *English (Second Language); English for Special Purposes; *Foreign Language Books; *Language Instruction; Language Research; Language Teachers; *Reading Comprehension; Science Education; *Second Language Learning; *Social Sciences; Systems Approach; Textbooks

ABSTRACT

This article describes the principles for planning an experimental teaching-learning system of the comprehension of English social and behavioral scientific texts at the university level. Results from a related evaluation study are also presented. The study was undertaken because of the unsatisfactory development of foreign language teaching and learning among the students of the University of Jyvaskyla whose main subjects were other than foreign languages. The foreign language teaching of the faculty is analyzed as a hierarchy of goal setting, methods planning, and evaluation. Four subsystems were identified: the foreign language examination board, the faculty of the social sciences department, the teacher, and the student. Three approaches to goal setting were considered: a trait-referenced approach, text-referenced approach, and needs-referenced approach. Instrumental objectives in reading comprehension of a foreign language text and measuring comprehension are discussed. Results of the first trial of the reading comprehension course are examined. The evaluation of the course focused directly on the student's subsystem (learning results, appropriateness of the needed initial level in English, and coordination of the course with the rest of the study program). (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED171138

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Dept. of English
Univ. Jyvaskyla

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

AN ENGLISH LANGUAGE TEACHING EXPERIMENT IN THE COMPREHENSION OF SCIENTIFIC TEXTS

RAIMO KONTTINEN
University of Jyväskylä

FL010329

This article describes the principles for planning an experimental teaching - learning system of the comprehension of English scientific texts at the university level in the social and behavioral sciences, and draws together the main results obtained from the related evaluation study. The planning principles could not be studied empirically, but the promising results of the experiment might justify their further development and study in their own right.

The practical impetus for the experiment was the unsatisfactory development of foreign language teaching and learning among the students of the University of Jyväskylä whose main subjects were other than foreign languages, hereafter called foreign language skills teaching and learning (FLST). The formal foreign language requirement of the university degree programs usually consisted of one or two so-called pro-exercitio courses leading up to a translation of an English text passage into Finnish. The objectives were rather obscure, the translation examination commonly used was criticized as a measuring instrument and the objectives which it represented were thought to be relatively irrelevant to other studies. The courses were taught by part-time teachers who naturally had little or no resources and no responsibility for the systematic development of teaching. Justified or not the criticism led to an even worse situation: students and faculties were inclined to decrease the foreign language requirements and a noticeable decrease in willingness to participate in the courses was attested.

The multitude and complexity of the problems soon made those involved realize the large extent of the problem area. It seemed impossible to change the situation by solving separately such single problems as student motivation, attitudes of faculties or examining boards, level of learning materials, etc. An improvement in one compo-

nent of the system (eg. increasing the length of the courses) would most probably have led to a worse situation in some other aspects (eg. decreased student motivation). The need for a systems-analytic frame of reference has been emphasized both in curriculum planning (Bloom 1966) and in educational planning in general (Hüfner and van Gendt 1971). It was also used in connection with the reform of the Finnish university degree programs (Konttinen et al. 1972) and it was considered useful in the present connection as well. Even a superficial analysis of the subsystems involved showed that a study of any single aspect of the foreign language teaching, eg. of new teaching methods or examining methods, would be of limited value only to the whole situation. It was decided, then, to try to plan an entire teaching-learning system and to study its effectiveness and some of its most central features empirically. It was felt that any research that avoided disturbing the teaching in operation would give results of academic interest only and that by creating changes, even if they were seen as disturbances, it was possible to gain valuable information.

The Foreign language skills teaching was conceived of as consisting of a hierarchy of Goals-Means-Evaluation chains. This is illustrated in Figure 1. For instance, a Faculty sets goals for its activities, chooses between possible degree programs, evaluates the success of its activities, and redefines and corrects the goals and procedures. The teaching-learning system can, however, be analysed from different points of view, eg. from the point of view of the Faculty, of a teacher, and of a student. These can be taken as subsystems, which are, at least partially, autonomous, but related to each other through actions, reactions and feedback. The differentiation of different subsystems becomes obvious, when the purpose and criteria of success of a foreign language course is studied from the point of view of a Faculty, of a teacher or of a student (Figure 2). It was thought that if the different subsystems or interest groups could be identified, at least some of the symptoms and difficulties of foreign language teaching could be diagnosed as results of conflicting goals or unsatisfactory means available or insufficient feedback.

Four subsystems were identified: the foreign language examination board (FLEB), the faculty in question (the Faculty of Social Sciences at the beginning of the experiment), the teacher, and the student. No other subsystems were considered in the planning of the experimental course,

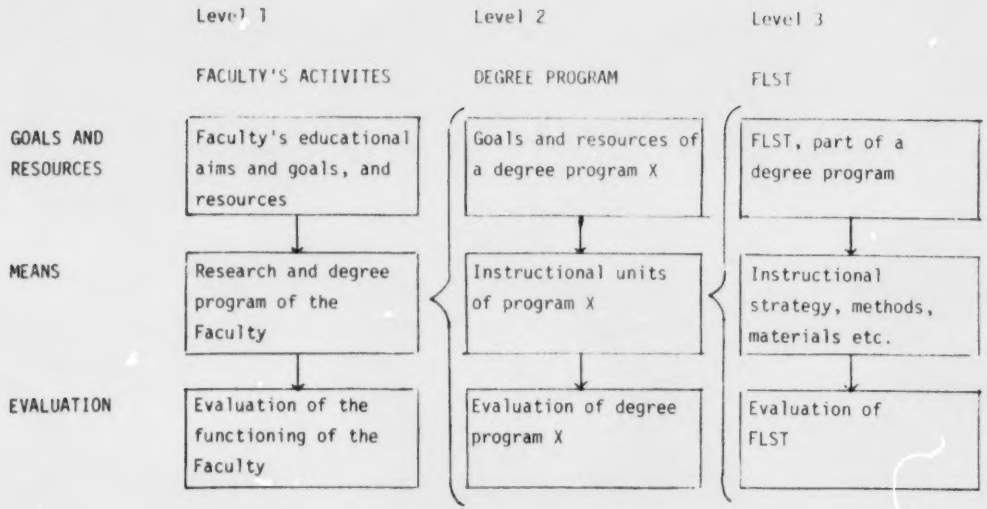

FIGURE 1. The foreign language teaching of a Faculty as a hierarchy of Goals - Means - Evaluation chains.

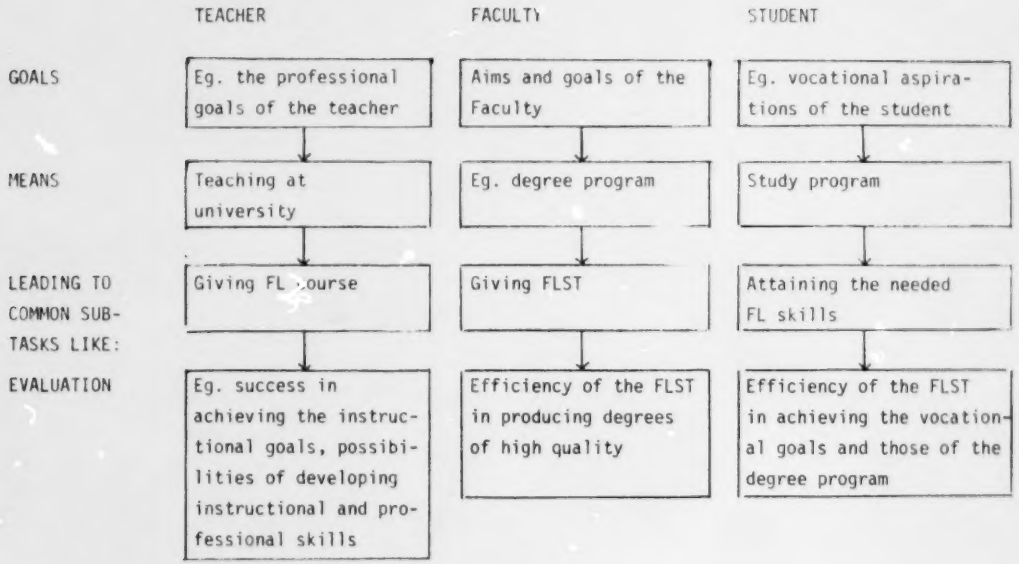
TEACHER
FACULTY
STUDENT
GOALS
Eg. the professional goals of the teacher
Aims and goals of the Faculty
Eg. vocational aspirations of the student
MEANS
Teaching at university
Eg. degree program
Study program
LEADING TO COMMON SUB-TASKS LIKE:
Giving FL course
Giving FLST
Attaining the needed FL skills
EVALUATION
Eg. success in achieving the instructional goals, possibilities of developing instructional and professional skills
Efficiency of the FLST in producing degrees of high quality
Efficiency of the FLST in achieving the vocational goals and those of the degree program

FIGURE 2. The foreign language skills course in the chain of Goals - Means - Evalution of a teacher, of a faculty, and of a student.

although they were highly relevant in the planning of the reform of foreign language skills teaching at the universities. Each one of the above mentioned subsystems carries out some kind of instructional planning. The FLEB defines the instructional objectives in the most concrete way, guarantees the proper level of requirements and has an impact on the work of the teachers. The faculty decides on the foreign language requirements in the degree program, but it cannot ignore the opinion of the FLEB or other language experts. In practice, the faculty has to accept the foreign language certificates given by the FLEB of the university. The teacher defines the teaching and learning objectives which he is aiming for in the course, but he is expected to prepare the students for the final examination given by the FLEB. The means, materials, time, etc. available to the teacher are rather limited and mostly controlled by other agents; in practice, the instructional planning and preparation remains rudimentary. The students also set goals, when choosing between training programs or study subjects and they probably accept the general goal of foreign language teaching, but they have little control over the way this goal is concretized, or over the means - other than self study - of achieving it.

The analysis showed that there was at least one point in the functioning of the different subsystems which was presumably a source of several problems: this was the goal conflict. The means and the feedback were not satisfactory either, but these difficulties were considered to be secondary to the goal conflict. An idea central to the planning of the experimental course was, then, the remedy of this conflict. It was only partially concerned with goal analysis and empirical study (to be described later) but, to a great extent with negotiation, the writing of proposals and persuasion. The problems were a clear identification and sharp definition of the needs of the faculty, an interpretation of the skill in understanding texts acceptable to the FLEB, and a goal formulation relevant, and seen to be relevant, from the point of view of the students and their studies and suitable for the instructional design of the course.

The necessity of defining goals would have been obvious even without an analysis of the whole system. There are, however, cases of a straightforward application of an educational technological model

(goal setting, methods planning, evaluation) by responsible decision makers (the faculty in this case) which have encountered difficulties in classroom practice or among the students. Curriculum planning is not only a multi-stage (goals, means, evaluation) but also a multi-phase process. In goal setting the latter means that the goals cannot be given by one of the subsystems if it is interlinked with the other subsystems. The goals have to be adjustable to those of the other subsystems and instrumental to them.

The systems-analytic thinking in the present context at least produced the different points of view on the goal setting problem, and it showed what is wrong in the goals of foreign language teaching and how the present goal setting is related to the problems and complaints of the different parties. The results were also used in making a proposal for the development of foreign language skills teaching at Finnish universities, which was prepared by a committee chaired by Kari Sajavaara (KKOP 1972) and which resulted in the foundation of the Language Centre of the Finnish universities.

Besides the systems-analytic thinking, a certain position was also taken up concerning the principles and conditions of learning. By means of studies comparing different teaching methods it has not been possible to find clear differences in their effectiveness, provided the objectives, motivation and initial level of the students and the other relevant factors have been properly controlled (Walker and Schaffarzinck 1974). The most prominent variable affecting learning is the amount of practice in the specific task observed and, consequently, the fact that a certain task was included in the curriculum at all.

It seems, then, to be difficult to produce better learning results by manipulating the teaching-learning system by means of the learning situation means of the variables within the learning experiment. It would be worth-while investigating the variables of defining the objectives, guaranteeing students' motivation, involving the teachers and so on, which are 'outside the actual learning situation'. It might be fruitful to discover, eg., under what conditions students are willing and able to work most effectively and benefit from the course, under what conditions the teachers are able to put their professional skills into full use, or what the learning activities are which bring the students closer to the

aims which the degree program designers had in mind.

In more general terms, it might be worth-while studying the co-ordination of the goals and activities of different subsystems, taking into account that each one of them is an autonomous, but not independent, body. Students, for example, have often been seen as objects of university teaching, and the problems of teaching are, first, how to build an efficient channel of information (teaching method) and, second, how to keep the students interested (motivation). Problems of this kind result in studying methods as general tools which transfer the information to the students under all circumstances. Students' motivation is studied independently of the students' aspirations in the hope of finding external incentives which students cannot resist. If the main problems are found in the learning situation, the work of the teacher and the faculty is seen as administrative practical details. They, too, could be taken as autonomous units defining their own goals and trying to achieve their aspirations and personal values through the teaching activities.

The main principles in the planning of the learning conditions were (1) to offer enough opportunity of learning the subskills not mastered by the student, and (2) to specify the objectives and to make the teaching-learning conditions compatible as far as possible with the goals and working conditions of students and also with those of the Faculty and the teachers. These broad principles led to an emphasis on three aspects in the planning of the course. First, an attempt was made to define the learning objectives accurately and communicate them to all persons and bodies involved. Secondly, a mastery learning strategy was accepted with the variation of study time (actual length of the course) according to the achievements of a student, and efforts were made to allocate the study time to the subskills which had not yet been mastered by the student on entering the course. Third, case was taken to integrate the foreign language teaching with other studies and schedules to obtain favourable practical conditions for foreign language learning. The latter aspect could not be properly observed since a systematic instructional planning of the degree programs along the lines of systems-analysis had not yet found its way into the faculties.

THE OBJECTIVES AND MEASUREMENT OF THE COMPREHENSION OF ENGLISH SCIENTIFIC TEXT

Three approaches to goal setting were considered. First, comprehension of a foreign language text, like other skills and abilities, can be considered as a trait of some stability over time and over different types of texts and measures. A university faculty, for example, may require from the students any kind of trait or mental faculty which exists and has been discovered by scientific research. This approach might be called *a trait-referenced approach*. In foreign language reading comprehension the emphasis is here on the comprehension trait or traits and not in what the student should do or shows himself capable of doing when comprehension is manifested and in what kind of situations he should show comprehension. The objectives are defined in terms of responses and not in terms of stimulus-response sets. The approach may, of course, reflect unanalytic goal setting, but it also avoids pragmatism and too narrow and mechanical or behavioristic goal setting. It may also be based on the idea that it is impossible to teach single stimulus-response bonds and that the broader classes of human behaviour that can be taught are the traits.

The trait-referenced approach has certain implications for the evaluation of learning results and teaching. Initially, it is necessary to find the trait or traits of comprehension of a foreign language text. There is plenty of research on the comprehension of mother tongue texts (eg. Davis 1972) and several traits have been discovered. When several traits are found in the domain, further specification of the learning objectives is needed, which necessitates some kind of weighting of the traits. The goal setter can choose between the traits or weight them, but he cannot define new traits by himself, because they are products of aptitudes and environmental effects. The measurements of the traits are, then, standardized tests of some kind, planned to measure as validly as possible the 'real' trait or trait combination.

The second approach could be termed *text-referenced*. An acceptable performance is defined with reference to the text. The student is asked to produce a new text, possibly in his mother tongue, which corresponds to the original. A student could, in principle, pass the examination if he could translate a sign warning of a danger of falling rocks, even though he might later be killed by one. The event itself is non-linguis-

tic in character and teaching somebody to avoid falling rocks is certainly not foreign language teaching. In this approach, a clear distinction is made between the knowledge of, and the skills in, the foreign language and their use and application. For evaluative purposes the tests can be translations or summaries, resumés, etc., but if more accurate decoding is required, it is more probable that translation will be chosen.

The third approach could be called *needs-reference*. Goal setting starts from the assessment of the training needs and other general educational aims. The needs are not necessarily, and not even usually, foreign language needs. As far as a faculty is concerned, they come from the goals of the degree programs and reflect desirable professional and scientific activities and readinesses. Foreign language goals come later, as objectives instrumental to the goal of the training programs. If the students are to be taught to avoid falling rocks in a foreign country, the objective of foreign language comprehension and the relevant course might be seen instrumental to this goal. In this case the student will pass an examination if, after reading a warning of falling rocks, he can avoid them, even though he cannot properly translate any one of the words or structures in the text, which is, of course, unlikely. In this approach any clear demarcation between language goals and language teaching, on the one hand, and other instruction, on the other, is not possible. Better relevance of the studies can be achieved provided that the goals of the degree program and other educational aims can be clearly analyzed. In evaluation, the approach easily leads to obtaining a work sample of some kind.

The three approaches are not mutually exclusive. They can all be exaggerated, which leads either to vague goal setting, word-for-word translation examinations, or a limited knowledge of the language. They can, however, be applied successfully, each in a different context and stage of goal setting. In the present context the needs-referenced approach was chosen for several reasons.

First, in the trait-referenced approach the traits are defined on the basis of the correlations between tests supposed to measure different aspects of the phenomenon. High correlations, however, do not necessarily imply a common construct behind the test performances.

Correlation coefficients can even be manipulated. A high correlation, for instance, between reading comprehension and auditory discrimination can be produced if some subjects receive exhaustive auditory-visual discrimination and comprehension training in these two skills, and other subjects are left without training. In the absence of experimental evidence for the processes which lie behind the test performances, eg. reading comprehension tests, it is too hazardous to rely on the traits found up to a certain point and give up specifying the needs of the training programs.

Secondly, the text-referenced approach was considered to offer too limited a basis for goal setting. As the aim was not to train interpreters or translators, the desired behavior of the students could not be defined in reference only to the text but also with reference to the use and application of the comprehension skill. The same limitation would be encountered when trying to define the universe of acceptable texts for comprehension tests.

The third approach was thought to offer a good starting point for goal setting in this context. Needs in training and education were felt to represent an appropriate level of conceptualization when the foreign language teaching problems were discussed either with the Faculty on the degree program level, with the students on the study plan level, or with the teachers on the level of instructional planning.

The needs could be derived from several sources. It was easy to find several good reasons for a need of English language reading comprehension. Arguments for other language skills could be found as easily, however. For instance, an oral skill is sought, for to prepare students for congresses, but it is not clear how many students will ever need a foreign language for this purpose. Even a knowledge of Latin has been suggested as a requirement because many scientific terms in Finnish come from Latin. In this project, the foreign language was considered to be instrumental and compatible with the goals and aims of the degree programs and this offered a natural basis for the needs.

Three sources for the foreign language needs were considered relevant. The first was, the goals of the degree programs. Unfortunately, they were changing with the reform of the degree programs. On three general goals, however, there seemed to be a sufficient consensus: the

degree programs will prepare students more explicitly than before for professional work in some field of society; they should prepare the students to apply scientific knowledge, thinking and research and not only practical skills in their work; there was a need for some kind of a general foreign language proficiency, but this was only secondary to the main goals of the programs.

The second source for the needs was the general nature of university studies in Finland. They are, of course, fixed, but some of their features seem to remain unchanged. There was a heavy reliance on foreign language textbooks from the very beginning and meagre foreign language contacts either in teaching situations or outside them. Thirdly, the foreign language skills of new students as compared with the goals of their studies could also be conceived as one source for the needs.

In studying the sources for the needs, it turned out that the need to acquire scientific knowledge and information through the medium of a foreign language, in English in the first place, was, and will be for some time, a major motivation for foreign language teaching and learning. At least two-thirds of the textbooks used at the Faculty of Social Sciences at the University of Jyväskylä were in English. Almost all the students know some English, but they have not had enough practice to be able to read fluently books in English containing scientific terminology and usage. Reading comprehension, as the only foreign language skill continuously in use, might also be enough to maintain the other skills.

The goals of the experimental foreign language course were derived from the assessment of the needs and they were formulated as follows. Taking into consideration the goals of the degree programs of the Faculty and other expressed aims, the probable study process, and the skills of the future student population, it is necessary for the students to comprehend general or elementary English texts in social and behavioral sciences from the beginning of their studies, to the extent that they are able to read a book in a week (i.e. read about 50 words per minute), and to obtain from the text all the factual information which could be used in a scholarly discussion and which could be obtained from the text without any previous knowledge

of the subject matter.

To improve the integration of foreign language teaching goals with the general goals of the degree programs, the training goals and aims have to be specified and clarified much further than is the case usually. It is then possible to decide whether any foreign language teaching is needed and, if if is needed, in what forms and in what ways it should be integrated with other studies. In this process there are, however, several problems. Some of them are practical, but further research is needed on many of them. The need for research results from the fact and experience that the process is not simple, logical and linear. More specific goals of a degree program cannot be derived logically from more general goals. The relationship is probably an interacting one in character, the specific goals somehow trying to respond to the needs stated in the more general goals. Furthermore, the process cannot be implemented directly or linearly. For instance, the actual effectiveness of foreign language instruction may lead to a redefinition of the more general goals, to a decrease of study time, an increase of language requirements, etc.

In connection with the present work, the following points in the goal specification chain were found to be especially in need of further research. The needs of the students' prospective professional work are, perhaps, the easiest to study. Examples are found in the reports by Roininen (1971, 1972a, 1972b) and Berggren (1975). The methodology of this kind of research is, however, in need of further study. Its usefulness for further goal specification depends on the needs descriptors or on the descriptive system of the use of the foreign language. A system of this kind is offered by Freihoff and Takala (1974). There is not, however, much experience of its usefulness in discovering the degree program needs. To what extent it can reveal training needs should be studied, not directly implying foreign language teaching, but relevant for the actual formulation of foreign language goals and the integration of foreign language teaching with other studies.

One line of research which is also highly relevant for foreign language goal setting at a university level is related to the problem of how students gain knowledge from texts. The problem has been accentuated in the Finnish context by the general aim of the new degree program to develop in the students a readiness to scientific and science-based activities in their specific fields of application. Because written materials will inevitably

be one of the most important sources of knowledge during and after pre-service studies, serious attention should be paid to this in foreign language teaching at the universities. The general problems include, eg., how students gain information from texts and how they accumulate it in structured knowledge; the possibility of differences arising from whether the subject matter is concrete, based on common experience, or abstract scientific matter, or whether the information attained has to be communicated further to others or applied. The studies on how students structure and recall texts will certainly have application in foreign language teaching, too, even though they are not directly related to it.

INSTRUMENTAL OBJECTIVES IN READING COMPREHENSION

The reading process was structured in terms of subskills or objectives supposed to be instrumental to the reading comprehension of a foreign language text. These were the general reading skill and two components of language, vocabulary and affixes, on the one hand, and structures, on the other. They were taught and evaluated, but not formally examined, because they were supposed to appear in the reading performance, if they really were instrumental to reading comprehension.

All the students were able to read, of course, but it was not so obvious that they were able to structure larger passages of text, to observe the function of a passage, and to adjust their reading style according to the nature of the text and reading purpose. These skills were taught in connection with some English reading materials of the experimental course and were measured with a questionnaire.

In choosing the vocabulary, an attempt was made to concentrate on words which are vital to the comprehension of elementary or general scientific texts and which are unknown to many students. In principle, words which most improved reading comprehension were accepted. In practice, about 80 to 90 per cent of the words to be encountered in the texts (i.e. running words) in typical scholarly writing in social and behavioral subjects were included. The remaining ten per cent of the words were supposed to be technical terms or words which could be understood with a sufficient knowledge of the most common affixes, whose meaning could be guessed. It was suggested that the teaching of

technical terminology should be included in the introductory courses of various departments. About 80 affixes were selected on the basis of the teachers' judgement and they were taught during the course. Intelligent guesswork was, if not taught, at least reinforced during the course.

The structures were also chosen by the judgement method, taking into consideration the previous knowledge of the students and the frequency of the structures in a scholarly writing. For the experimental course, about 30 structures were chosen for explanation and practice during the course.

MEASURING THE COMPREHENSION OF FOREIGN LANGUAGE SCIENTIFIC TEXTS

The available means for measuring were mostly translation tests, cloze tests, or multiple choice tests for the comprehension of text passages. No available tests seemed to suit as such for the evaluation of the attainment of the present objectives. The means to be used to gauge the general goal of the present course were to fulfil a number of requirements:

- the test should probe the factual information in the text, not words and structures as such;
- the test should not measure an ability to match English to a Finnish word or translation but, rather, measure the skill to transform the content into other words, perhaps into the mother tongue;
- the test should not require an active use of English;
- it should not require any prior knowledge of the subject matter;
- it should measure the understanding of the main ideas, facts and argumentation of the text;
- it should not, however, depend on the reasoning ability of the students; or
- it should not measure the ability to memorize the text; and
- it should not measure the student's perceptual ability covered by items, like "On what line is the word abc?".

One measuring procedure, and perhaps the ideal one, is to let subject-matter experts examine the comprehension of the text by the students in a thorough discussion of the text. In fact, the measuring instruments were developed to stimulate such a situation. It was decided to give the examinees several (six) texts, each about 500 words in length. This was thought to be a sufficient text sample from the type of the social and

behavioral science texts which the students were to encounter in their studies. On the basis of each text, some ten multiple-choice items were written in Finnish in accordance with the above requirements. Many trial texts and items were rejected by the team and subject-matter experts as improbable in a learned discussion of the content of the texts, as obscure, or for other similar reasons.

Formal item analyses were also carried out (difficulty indexes and item test correlations), but the results were used only as cues for inappropriateness of an item. Other item formats were considered and tried, eg., translation of short passages, but the multiple choice format was accepted for practical reasons. Skimming and resumé writing might also have been included, but they were also rejected for practical reasons.

The test took about two hours. The exact time was decided by counting one minute for each question and allowing one minute of reading time pro 50 words. This time turned out to be sufficient for about 95 per cent of the examinees. This result also showed that about 95 per cent of the students can fulfil the reading speed objective.

To pass the examination, the student was supposed to understand almost as much (about 90 per cent) of the text as of a similar Finnish text. The answers to the remaining ten per cent of the questions were supposed to become evident after the students were familiarized with the meaning of the technical terms. A number of tests were based on English books which were available in Finnish translation. One pair of comparable English and Finnish reading comprehension tests was administered to the students. The average score in the Finnish version was about 41 out of 60 questions. Since a score of 40 correct items out of 60 was found to be comparable with the standards of FLEB, it was set as the minimum passing score.

It was also afterwards checked that it was not possible to pass the examination with a sufficient amount of general subject matter knowledge. The questions of two tests without the texts were given to two groups of advanced students. The average scores were around 30, which is better than scores usually obtained by mere guessing but clearly less than the minimum passing score. It was concluded that the tests did not measure a general knowledge of the subject matter, but rather

the understanding of more specific information given by the texts.

A number of other criteria for the pass level were also considered. The reading skill should be high enough to make the study of English books comfortable and to encourage learning from written English materials without relying on any translation or summary services on the campus. The reading skill should also be good enough to maintain the students' other English language skills.

The reading comprehension tests were of sufficiently high reliability. Split-half coefficients varied between .9 and .7, depending on the heterogeneity of the examinee group. The coefficients are not, however, the most appropriate in this context for revealing the general application of the test results. The aim was not to produce tests to discriminate between students or to make comparisons between them but rather to compare the student with the learning objectives, and then somehow to equate the tests to give comparable results, which is normally the case with common standardized tests or norm-referenced tests (Millman 1974). In this study the aim was to design criterion-referenced tests (Millman 1974) which would be comparable due to the very procedure by which they were produced. Each test should be appropriate and as good as the others, if it is to be a good sample of the behavior stated in the learning objectives, and the results which it gives should give a good approximation of those obtained by giving the examinee all possible parallel tests.

To study the generalizability (Cronbach et al. 1972) of the score of one test to the corpus of all possible tests, five versions were administered to 55 students each. The results showed that the actual pass level or the difficulty of a test may shift from test to test, even though the instructions given to the item writers were supposed to be accurate. But, even with this source of error, the reliability or, rather, the generalizability of the test scores was reasonable. The standard error was 4.6 and the alfa reliability coefficient .7 on average. This means that in 68 per cent of the tests and measurements a student's true score will be in the range of plus/minus 4.6 points around the score given by the test. Furthermore, it turned out that it would be profitable to increase the text sample to ten or even to 20, rather than to increase the number of items per text to obtain a more reliable score. The effect of increasing the number of text passages and making the problems of the tests more comparable would mean a decrease of about .8 in the standard error.

The results of the reliability and generalizability studies show that the tests give a quite accurate picture of the examinee's reading skill and that the different versions of the test give sufficiently comparable results. In the studies with a translation or other examination based on one text passage the same level of comparison from test to test or generalizability of the results to the corpus of texts and tests could not be reached.

TABLE 1. Success of the students on the first application of the experimental course in the Faculty of Social Sciences (FSS) and the Faculty of Physical Education (FPE).

Student group	FSS-71	FSS-72	FPE-72
1. Passed the Exam earlier	6	23	15
2. Did not show up	13		
3. Attended, but gave up	18	8	39
4. Attended regularly, failed	3	1	3
5. Passed the Exam	60	68	43
Total per cents	100	101	100
New students total	222	267	88
Passed of groups 3-5	73	88	51

AN EXPERIMENT WITH THE FIRST TRIAL READING COMPREHENSION COURSE AND MAIN RESULTS RELATED TO ITS EFFECTIVENESS

The material for the first trial course was designed, according to the principles described earlier, by Eva May. It included material for teaching different reading styles, affixes, vocabulary and structures. New versions of the material are available from the Language Centre at the University of Jyväskylä.

The trial version of the course was administered to all new students at the Faculty of Social Sciences (FSS) in the autumn terms 1971 and 1972 and also for the students from the Faculty of Physical Education (FPE) at the university of Jyväskylä in 1972. The students were required to sit an examination at the beginning of the course and then once a month until they passed the examination. At the first examination several other tests were given to the students, but only

the tests of vocabulary, affixes, and structures were given at the monthly examinations. (For more details of the data collection see Konttinen and Sikanen 1976; in the present paper only the main results will be given.)

The first question of interest among the results is the question of whether the students learned reading comprehension during the course. The answer is positive. This is not, however, a very interesting result as such, because learning is a common experience of all teachers and students and, besides, the students were not passed and were urged to stay on, until they reached the objectives. A question which is more interesting from the point of view of the general ideas behind the planning of the course, is whether the students accepted the objectives and were willing to work and able to achieve the objectives under the circumstances created.

The teaching-learning system turned out to be effective in the sense that about 95 per cent of the students who persisted with the course, to the end of the term if necessary, achieved the reading comprehension objective (see Table 1.). This could be taken as a piece of evidence for the success of the experimental course. If, however, the whole teaching-learning system (program planning and the Faculty, instructional planning and the teachers, studies planning and the students) is taken into account, the situation looks quite different. From the point of view of instructional planning or the planning and execution of the experimental course, the results were promising. In this respect, the fact that many students were not willing or able to attain the objective can be regarded as being due to the decisions of the students, to their previous training, or to the Faculty, but not to instructional planning and to people involved in it.

From the point of view of both the Faculty and the student the course was not as successful. About 65-70 per cent of the new students at the Faculty of Social Sciences (FSS) who had not passed the examination earlier attained the objective during the autumn term. At the Faculty of Physical Education (FPE) the corresponding figure was only about 45 per cent. If both the students who had fulfilled the requirement earlier and those who could not be contacted at all, probably because they were not full-time students, are excluded from the failure cases, the success rate of the teaching-learning system becomes somewhat better (73, 88, and 51 per cent, see Table 1). The experimental course presumably compared quite well with the former translation-type course. The second, improved, version of the

course (particularly in the FSS, for which it was specially designed), especially, was so successful that it could be implemented with minor modifications. The course cannot, however, be considered flawless such as it is and further revisions are necessary.

The results of the experiment support the view that in the development of a teaching-learning system the teaching method or the teaching materials may not be the bottle-neck, but the source of the problems will have to be sought in some other factors which are outside the actual teaching and learning situation. The results of the diagnostic tests (concerning vocabulary, affixes, and structures) and the questionnaire revealed three such factors:

(1) About 15-20 per cent of the students obviously did not have the command of the most basic vocabulary and structures.

(2) About 15-20 per cent of the students at the FSS had a very crowded timetable and could not find enough time to attend the course properly; at the FPE, (giving each student about 40 hours in a week obligatory courses) the number of students who reported the same reason was considerably larger.

(3) The content and texts of the course designed for students of social and behavioral sciences were seen as irrelevant and unmotivating by the students of physical education.

The three main reasons for the failures are easy to understand and, in principle, even easy to remedy: what is needed is a preparatory course based on the diagnostic test results, better co-ordination and scheduling of the compulsory courses, and adaptation of the course materials to the broad field of studies. It would also be easy to put the corrective measures in practice, if some person or organ had a complete command of all the relevant subsystems. They cannot and should not, however, be changed directly on the basis of research results. What research can do is to give a better conceptualization of a multi-stage and multi-phase process of educational planning and find models and practical ways to manage such a complex process. In this connection, research on educational innovation (eg., Sashkin, Morris and Horst 1973) could add valuable experience to a more theoretical system-analysis-oriented research.

One question which is relevant to the adaptation of the mastery

learning strategy is related to learning speed or learning ability. If the individual speed of learning is seen as an essential determiner of achievement, teaching should be individualized or differentiated on the basis of this trait. As learning speed had often been equated with intelligence, or perhaps in contexts of this kind with verbal ability or foreign language aptitude, individualization would mean ability grouping of some kind. But, if the amount which an individual has to learn to attain the objective is seen as a more important determinant of achievement, it may be more useful to let the study time vary and direct the efforts towards the subskills which the individual will have to attain to reach the objective.

To investigate the existence of a learning aptitude, the gains in learning of the vocabulary, affixes, structures, translation, and reading comprehension skills during the first month were correlated with each other. The correlation coefficients were low, which does not support the view of the importance of the learning speed on achievements.

Another way to solve the problem was to try to differentiate the supposedly quick learners who passed the examination in the first test. A discrimination analysis was made of the variables which described the students' achievements in English and their aptitudes in the groups which passed the examination in September, October, November, and December, and in the group of those who never passed the test. The most prominent discriminant function differentiated, in fact, the groups in this order. What was interesting in the discriminator was that it indicated a student's proficiency in English but not his verbal or reasoning ability or memory. A student achieved the objective faster or slower depending on how far he was from the comprehension goal, and not depending on his aptitudes.

The results of both studies support the decision to adapt the mastery learning strategy to this course. At least among the university students who try to improve their foreign language reading skill, the results do not support the view that the most able students are the first to learn the skill and that an ability grouping of some kind would be profitable. Aptitude and good learning abilities are certainly not unimportant, but the level and nature of previous training and the student's own motivation seem to easily overshadow the differences in the aptitudes. The results could not be interpreted to support undifferentiated teaching either. Individualization can be, and was supposed to be useful in the present

experiment, but it should aim at providing training in such parts of the curriculum as are not yet mastered, until the goal is attained. A comprehensive grouping based on some general ability or aptitude measure would not, then, be as profitable as a specific differentation of the instruction based on the specific difficulties of learning.

RESULTS RELATED TO THE MEASUREMENT OF THE ATTAINMENT OF THE OBJECTIVES

Other results of the studies which were related to the first two trials with the experimental course illuminate the content of the reading comprehension test and the role of the instrumental skills in reading performance. A factor analysis was performed on the reading test, and on the tests of vocabulary, affixes, structures, translation, reading speed, listening comprehension, writing, linguistic analysis, reasoning ability, and verbal ability, and on the school marks and the questionnaire of study and reading habits. All the tests were administered at the beginning of the first experimental course (1971) to all the new students in the Faculty of Social Sciences.

The reading comprehension test which was developed to measure the attainment of the reading objective turned out to be quite a good measure of the General English Achievement, a factor containing both the tests of the components of language (vocabulary, affixes and structures) and the integrated skills of reading, listening, translating, and writing. To get a better idea of the nature of the reading test, it is easier to describe what it did not measure. It did not depend on the factors of Response Speed (in reading tests), Ability to Analyze and Learn an Artificial Language, Verbal Ability, or on the factors of study and reading habits. It did correlate slightly with the Reading Speed in English, which is in accordance with the learning objective.

The principles of the reading test design were derived from the necessary uses of the skill and aimed at measuring the skill of gaining information from a text, in the first place, and not linguistic skills, or ingenuity. In the beginning it was not at all clear whether the test would measure verbal reasoning as most mother tongue reading tests do, or general reading skill, which is independent of the language, and certain doubts were expressed whether it could be taken as a foreign language proficiency test at all unless some language component tests

were added to it. The tests developed seem to fulfil their purpose quite well.

The interrelationship between reading comprehension and vocabulary, affixes, and structures, was studied further, to find out to what extent reading skill improves with an improvement in these instrumental skills. An inspection of the scatter diagrams of the reading test with the tests of vocabulary, affixes and structures could not reveal any clear thresholds. Reading skill seemed to increase quite steadily with a better knowledge of vocabulary, affixes and structures. It might be possible to find such thresholds, but the general nature of the subskill tests may have prevented their discovery.

If, however, the minimum pass score in the reading test can be taken as the limit where comprehension clearly begins, it is meaningful to study what the students who have achieved this limit know. This was accomplished by estimating the regressions of the reading test on the tests of the vocabulary, affixes and structures. The result was that a student already has a fifty-fifty chance of attaining an acceptable score (40 points out of 60) with a vocabulary of about 2200 words, which means 50-55 per cent of the vocabulary included in the course and containing at least 1000 most common words. A student with a larger vocabulary would probably have a better reading skill, but even with a vocabulary of only a little more than 2000 words the vocabulary as such should no longer be any obstacle in comprehending learned social and behavioral science texts.

The results concerning the role of the vocabulary lead to the question of the nature of those 2000 words. They may be a certain set of approximately 4200 words included in the glossary of the course. It may also be immaterial which 2000 words are known provided that they come from the present glossary of 4000 words of from any other relevant set of words. The answers could not be found in the data of the present study, and further research with a proper framework is needed to give information about the selection of vocabulary for similar purposes.

About 75 per cent of the affixes seem to be necessary for the attainment of the required comprehension level. The affixes were chosen by means of expert judgement, but the figure is ambiguous. It may imply that the right affixes were chosen and that about 3/4 of them are very important to comprehension. It may also imply that 75 per cent of any affixes, even

those not included in the present collection, are very helpful to the reader. Relying on the expertise of the teacher who chose the affixes the first interpretation of the result seems more plausible.

About 80 per cent of the structures chosen for special attention on the course turned out to be familiar even to the students lying at the 40 point limit. In selecting the structures it was supposed that all the more frequent structures were known to a majority of the students. The figure shows that the reading objective sets very high requirements on the knowledge of structures.

FURTHER DEVELOPMENT OF THE FOREIGN LANGUAGE COURSE

The evaluation of the foreign language course focused directly on the student's subsystem (learning results, appropriateness of the needed initial level in English, co-ordination of the course with the rest of the study program). The systems-analytic frame would require a similar evaluation from the point of view of the FLEB and of the faculties in question (the relation of the foreign language achievements to the content and goals of the degree programs, costs of the implementation and running of the new course) and from the point of view of the teachers (possibilities for flexible management of learning, time pressure etc.). These points were not studied systematically. The experimental courses gave, however, sufficient informal evidence to continue efforts along similar lines. At present the same design principles for the organization of the FLST and for the material have been applied to the reading comprehension course for the humanities and medicine. All the material contains booklets for Reading Comprehension Exercises, Structures, Affixes, Wordlists, and Teacher's Book and Quizzes, distributed through the Language Centre, University of Jyväskylä (SF-40100 Jyväskylä 10, Finland).

BIBLIOGRAPHY

Berggren, Ola. 1975. Teollisuuden palveluksessa toimivien insinöörien kielitaidon tarve ja insinöörikunnan käsitykset teknillisissä korkeakouluissa annettavan kieltenopetuksen sisällöstä. Lappeenrannan teknillisen korkeakoulun Kielten opetuksen keskuksen julkaisuja 1/1975.

Bloom, B. S. 1966. The Role of the Educational Sciences in Curriculum Development, International Journal of Educational Sciences 1, 5-16.

Cronbach, L. J., G. C. Gleser. H. Nanda, and N. Rajaratnam. 1972. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: Wiley.

Davis, F.B. 1972. Psychometric Research on Comprehension in Reading, Reading Research Quartely 7, 628-678.

Freihoff, R. and S. Takala. 1974. Kielenkäyttötilanteiden erittelyyn perustuva kieltenopetuksen tavoitekuvausjärjestelmä - The Systematic Description of Language Teaching Objectives Based on the Specification of Language Use Situations, Reports from the Language Centre No. 2, University of Jyväskylä.

Hüfner, K. and R. van Gendt. 1971. Towards a New Theoretical Framework in Educational Planning, Development and Change 3 (2), 18-27.

KKOP - Korkeakoulujen kielten opetuspalvelun kehittäminen - Development of the Foreign Language Skills Teaching at the Universities, 1972. Publications of the Department for Higher Education No. 1, Ministry of Education.

Konttinen, R., A. Nuutinen, M. Panhelainen, and M. Rantaniemi. 1972. Korkeakouluopetuksen suunnittelu ja kehittäminen - Planning and Development of University Teaching, Publications of the Department for Higher Education No. 2, Ministry of Education.

Millman, J. 1974. Criterion-Referenced Measurement, in W. J. Popham (ed.), Evaluation in Education: Current Applications, Berkeley, California: McCutchan.

Roininen, V. 1971. Elinkeinoelämän vieraiden kielten määrällinen ja laadullinen tarve Varsinais-Suomessa. Turun yliopiston Sosiologian laitoksen julkaisuja 38/1971.

Roininen, V. 1972a. Teollisuuden, kaupan ja eräiden palveluelinkoinojen ammateissa tarvittava vieraiden kielten taito. Turun yliopiston Sosiologian laitoksen julkaisuja 46/1972.

Roininen, V. 1972b. Teollisuuden, kaupan ja eräiden palveluelinkeinojen vieraiden kielten nykyinen ja arvioitu vuoden 1980 määrällinen ja laadullinen tarve Suomessa. Turun yliopiston Sosiologian laitoksen julkaisuja 50/1972.

Sashkin, M., W.C. Morris and L. Horst. 1973. A Comparison of Social and Organizational Change Models: Information Flow and Data Use Process, Psychological Review 80, 510-526.

Walker, D. F. and J. Schaffarzick. 1974. Comparing curricula, Review of Educational Research 44, 83-111.